

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局(43) 国际公布日:
2001年8月23日(23.08.01)

PCT

(10) 国际公布号:
WO 01/60855 A1

- (51) 国际分类号⁷: C07K 14/47, C12N 15/12, 15/63 (74) 代理人: 上海专利商标事务所(SHANGHAI PATENT & TRADEMARK LAW OFFICE); 中国上海市桂平路435号, Shanghai 200233 (CN).
- (21) 国际申请号: PCT/CN01/00121 (81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW
- (22) 国际申请日: 2001年2月12日(12.02.01)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权: 00111697.5 2000年2月17日(17.02.00) CN (84) 指定国(地区): ARIPO专利(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG)
- (71) 申请人(对除美国以外的所有指定国): 上海市肿瘤研究所(SHANGHAI CANCER INSTITUTE) [CN/CN]; 中国上海市斜土路 2200 弄 25 号, Shanghai 200032 (CN).
- (72) 发明人;及
- (75) 发明人/申请人(仅对美国): 顾健人(GU, Jianren) [CN/CN]; 中国上海市斜土路 2200 弄 25 号, Shanghai 200032 (CN)。杨胜利(YANG, Shengli) [CN/CN]; 中国上海市漕宝路 500 号, Shanghai 200233 (CN)。
- 本国际公布:
— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期 PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: A NOVEL HUMAN CELL CYCLE CONTROL-RELATED PROTEIN AND A SEQUENCE ENCODING THE SAME

(54) 发明名称: 新的人细胞周期控制相关蛋白及其编码序列

(57) Abstract: The present invention discloses a novel human tumor-suppressing protein and a polynucleotide encoding the same, as well as a method of producing the polypeptide by recombinant technique. The present invention also discloses methods of using the polypeptide in treatment of various disease, such as tumor or the like. The present invention also discloses an antagonist against the polypeptide and the therapeutic use of the same. Also disclosed is the use of such polynucleotide encoding the novel human tumor-suppressing protein.

(57) 摘要

本发明公开了一类新的具有抑癌功能的人蛋白, 编码此多肽的多核苷酸和经重组技术产生该多肽的方法。本发明还公开了此多肽用于治疗多种疾病如癌症等的方法。本发明还公开了抗此多肽的拮抗剂及其治疗作用。本发明还公开了编码这类新的具有抑癌功能的人蛋白的多核苷酸的用途。

(22-2586位)或全长序列(1-2659位)。

在本发明的第三方面,提供了含有上述多核苷酸的载体,以及被该载体转化或转导的宿主细胞或者被上述多核苷酸直接转化或转导的宿主细胞。

在本发明的第四方面,提供了制备具有 CLG 蛋白活性的多肽的制备方法,该方法包含: (a)在适合表达 CLG 蛋白的条件下,培养上述被转化或转导的宿主细胞; (b)从培养物中分离出具有 CLG 蛋白活性的多肽。

在本发明的第五方面,提供了与上述的 CLG 蛋白多肽特异性结合的抗体。还提供了可用于检测的核酸分子,它含有上述的多核苷酸中连续的 10-800 个核苷酸。

在本发明的第六方面,提供了一种药物组合物,它含有安全有效量的本发明的 CLG 蛋白多肽以及药理学上可接受的载体。这些药物组合物可治疗癌症以及细胞异常增殖等病症。

本发明其它方面由于本文的公开内容,对本领域的技术人员而言是显而易见的。

图 1 是 CLG 的 RNA 表达谱。

图 2 显示了 CLG 全长 cDNA 对人肝癌细胞 SMMC-7721 集落形成的抑制作用。其中图 2A 为载体对照(pCMV-Script),集落数为 56 个;图 2B 为 CLG(pCMV-Script/CLG),集落数仅为 3 个。

图 3 是用 CLG 转染的 SMMC-7721 细胞接种裸鼠后形成的肿瘤组织的石蜡切片照片。图 3A 为坏死灶内,可见较多阳性凋亡细胞;图 3B 为未坏死区,可见散在的凋亡细胞。

图 4 显示了转染 CLG 基因后的“梯形”DNA(DNA ladder)电泳图。其中各泳道为: 1. 转染空载体; 2. 转染 p53; 3. 转染 CLG; 4. 未转染对照点; 5. 分子量标准。

图 5 CLG 蛋白表达的 SDS-PAGE 电泳图。其中,各泳道是: 1. 蛋白分子量标准; 2. pET32a-CLG 菌体超声上清; 3. pET32a 菌体超声上清; 4. pET32a-CLG 菌体超声沉淀; 5. pET32a 菌体超声沉淀。

发明详述

本发明采用大规模 cDNA 克隆转染癌细胞,在获得具有抑癌作用的基础上,经测序证明为新的基因,进一步得到全长 cDNA 克隆。DNA 转染试验证明,本发明的具有抑癌功能的蛋白对癌细胞(肝癌细胞)具有抑制克隆形成的作用,其抑制率 $\geq 50\%$ 。

在一个实例中,从人胎盘 cDNA 文库中分离到一个与线虫、果蝇和酵母的细胞周期调控蛋白高度同源的新基因,它与线虫的细胞周期调控蛋白具有 66%(557/837)同源性,与果蝇的细胞周期调控蛋白(crooked neck, crn)有 41%(157/375)同源性,与酵母的细胞周期控制蛋白有 37%(228/599)同源性。从生物信息学可以判定人 CLG(crn-like gene; 原序号为 PP3898, GenBank 登录号 AF258567, 登录日期 2000 年 4 月 24 日)是一个新的人细胞周期相关基因。经过初步功能研究发现,CLG 可在体外抑制人肝癌细胞 SMMC-7721 的生长;肝癌细胞 SMMC-7721 裸鼠移植瘤组织的原位凋亡检测和 DNA 电泳分析表明,CLG 基因转染 SMMC-7721 细胞后能诱导肿瘤细胞凋亡,并抑制肿瘤细胞的生长。

在本文中,“细胞周期控制相关基因蛋白”、“CLG 蛋白”、和“PP3898 蛋白”可

更佳地至少 80%相同性的多核苷酸。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中，“严格条件”是指：(1)在较低离子强度和较高温度下的杂交和洗脱，如 $0.2\times\text{SSC}$ ，0.1%SDS，60℃；或(2)杂交时加有变性剂，如 50%(v/v) 甲酰胺，0.1%小牛血清/0.1% Ficoll，42℃等；或(3)仅在两条序列之间的相同性至少在 95%以上，更好是 97%以上时才发生杂交。并且，可杂交的多核苷酸编码的多肽与 SEQ ID NO: 2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与上述的序列杂交的核酸片段。如本文所用，“核酸片段”的长度至少含 15 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码 CLG 蛋白的多聚核苷酸。

本发明的 DNA 序列能用几种方法获得。例如，用本领域熟知的杂交技术分离 DNA。这些技术包括但不局限于：1)用探针与基因组或 cDNA 文库杂交以检出同源性核苷酸序列，和 2)表达文库的抗体筛选以检出具有共同结构特征的克隆的 DNA 片段。

编码 CLG 蛋白的特异 DNA 片段序列产生也能用下列方法获得：1)从基因组 DNA 分离双链 DNA 序列；2)化学合成 DNA 序列以获得所需多肽的双链 DNA。

上述提到的方法中，分离基因组 DNA 最不常用。当需要的多肽产物的整个氨基酸序列已知时，DNA 序列的直接化学合成是经常选用的方法。如果所需的氨基酸的整个序列不清楚时，DNA 序列的直接化学合成是不可能的，选用的方法是 cDNA 序列的分离。分离感兴趣的 cDNA 的标准方法是从高表达该基因的供体细胞分离 mRNA 并进行逆转录，形成质粒或噬菌体 cDNA 文库。提取 mRNA 的方法已有多种成熟的技术，试剂盒也可从商业途径获得(Qiagene)。而构建 cDNA 文库也是通常的方法(Sambrook, et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, New York, 1989)。还可得到商业供应的 cDNA 文库，如 Clontech 公司的不同 cDNA 文库。当结合使用聚合酶反应技术时，即使极少的表达产物也能克隆。

可用常规方法从这些 cDNA 文库中筛选本发明的基因。这些方法包括(但不限于)：(1)DNA-DNA 或 DNA-RNA 杂交；(2)标志基因的功能出现或丧失；(3)测定 CLG 蛋白的转录本的水平；(4)通过免疫学技术或测定生物学活性，来检测基因表达的蛋白产物。上述方法可单用，也可多种方法联合应用。

在第(1)种方法中，杂交所用的探针是与本发明的多核苷酸的任何一部分同源，其长度至少 15 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸。此外，探针的长度通常在 2kb 之内，更佳地为 1kb 之内。此处所用的探针通常是在本发明的基因 DNA 序列信息的基础上化学合成的 DNA 序列。本发明的基因本身或者片段当然可以用作探针。DNA 探针的标记可用放射性同位素，荧光素或酶(如碱性磷酸酶)等。

在第(4)种方法中，检测 CLG 蛋白基因表达的蛋白产物可用免疫学技术如 Western 印迹法，放射免疫沉淀法，酶联免疫吸附法(ELISA)等。

应用 PCR 技术扩增 DNA/RNA 的方法(Saiki, et al. Science 1985;230:1350-1354)被优选用于获得本发明的基因。特别是很难从文库中得到全长的 cDNA 时，可优选使用 RACE 法(RACE-cDNA 末端快速扩增法)，用于 PCR 的引物可根据本文所公开的本发明的序列信

使转录得到增强。增强子是 DNA 的顺式作用因子，通常大约有 10 到 300 个碱基对，作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

- 5 用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时，能吸收 DNA 的感受态细胞可在指数生长期后收获，用 CaCl_2 法处理，所用的步骤在本领域众所周知。可供选择的是用 MgCl_2 。如果需要，转化也可用电穿孔的方法进行。当宿主是真核生物，可选用如下的 DNA 转染方法：磷酸钙共沉淀法，常规机械方法如显微注射、电穿孔、脂质体包装等。

- 10 获得的转化子可以用常规方法培养，表达本发明的基因所编码的多肽。根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法（如温度转换或化学诱导）诱导选择的启动子，将细胞再培养一段时间。

- 15 在上面的方法中的重组多肽可包被于细胞内、细胞外或在细胞膜上表达或分泌到细胞外。如果需要，可利用其物理的、化学的和它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但并不限于：常规的复性处理、用蛋白沉淀剂处理（盐析方法）、离心、渗透破菌、超处理、超离心、分子筛层析（凝胶过滤）、吸附层析、离子交换层析、高效液相层析（HPLC）和其它各种液相层析技术及这些方法的结合。

- 20 重组的人 CLG 蛋白或多肽有多方面的用途。这些用途包括（但不限于）：直接做为药物治疗 CLG 蛋白功能低下或丧失所致的疾病，和用于筛选促进或对抗 CLG 蛋白功能的抗体、多肽或其它配体。例如，抗体可用于激活或抑制人 CLG 蛋白的功能。用表达的重组人 CLG 蛋白筛选多肽库可用于寻找有治疗价值的能抑制或刺激人 CLG 蛋白功能的多肽分子。

- 25 本发明也提供了筛选药物以鉴定提高（激动剂）或阻遏（拮抗剂）人 CLG 蛋白的药剂的方法。激动剂提高人 CLG 蛋白刺激细胞增殖等生物功能，而拮抗剂阻止和治疗与细胞过度增殖有关的紊乱如各种癌症。例如，能在药物的存在下，将哺乳动物细胞或表达人 CLG 蛋白的膜制剂与标记的人 CLG 蛋白一起培养。然后测定药物提高或阻遏此相互作用的能力。

- 30 人 CLG 蛋白的拮抗剂包括筛选出的抗体、化合物、受体缺失物和类似物等。人 CLG 蛋白的拮抗剂可以与人 CLG 蛋白结合并消除其功能，或是抑制人 CLG 蛋白的产生，或是与多肽的活性位点结合使多肽不能发挥生物学功能。人 CLG 蛋白的拮抗剂可用于治疗用途。

- 35 在筛选作为拮抗剂的化合物时，可以将 CLG 蛋白加入生物分析测定中，通过测定化合物影响 CLG 蛋白和其受体之间的相互作用来确定化合物是否是拮抗剂。用上述筛选化合物的同样方法，可以筛选出起拮抗剂作用的受体缺失物和类似物。

本发明的多肽可直接用于疾病治疗，例如，各种恶性肿瘤、和细胞异常增殖等。

本发明的多肽，及其片段、衍生物、类似物或它们的细胞可以用来作为抗原以生产抗体。这些抗体可以是多克隆或单克隆抗体。多克隆抗体可以通过将此多肽直接注射动

本发明中的抗体可用于治疗或预防与人 CLG 蛋白相关的疾病。给予适当剂量的抗体可以刺激或阻断人 CLG 蛋白的产生或活性。

5 抗体也可用于设计针对体内某一特殊部位的免疫毒素。如人 CLG 蛋白高亲和性的单克隆抗体可与细菌或植物毒素(如白喉毒素,蓖麻蛋白,红豆碱等)共价结合。一种通常的方法是用巯基交联剂如 SPDP,攻击抗体的氨基,通过二硫键的交换,将毒素结合于抗体上,这种杂交抗体可用于杀灭人 CLG 蛋白阳性的细胞。

多克隆抗体的生产可用人 CLG 蛋白或多肽免疫动物,如家兔,小鼠,大鼠等。多种佐剂可用于增强免疫反应,包括但不限于弗氏佐剂等。

10 人 CLG 蛋白单克隆抗体可用杂交瘤技术生产(Kohler and Milstein. Nature, 1975, 256:495-497)。将人恒定区和非人源的可变区结合的嵌合抗体可用已有的技术生产(Morrison et al, PNAS, 1985, 81:6851)。而已有的生产单链抗体的技术(U.S. Pat No. 4946778)也可用于生产抗人 CLG 蛋白的单链抗体。

能与 CLG 蛋白结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时,必须对人 CLG 蛋白分子进行标记。

15 本发明还涉及定量和定位检测人 CLG 蛋白水平的诊断试验方法。这些试验是本领域所熟知的,且包括 FISH 测定和放射免疫测定。试验中所检测的人 CLG 蛋白水平,可以用作解释人 CLG 蛋白在各种疾病中的重要性和用于诊断 CLG 蛋白起作用的疾病。

CLG 蛋白的多聚核苷酸可用于 CLG 蛋白相关疾病的诊断和治疗。在诊断方面,CLG 蛋白的多聚核苷酸可用于检测 CLG 蛋白的表达与否或在疾病状态下 CLG 蛋白的异常表达。
20 如 CLG 蛋白 DNA 序列可用于对活检标本的杂交以判断 CLG 蛋白的表达异常。杂交技术包括 Southern 印迹法, Northern 印迹法、原位杂交等。这些技术方法都是公开的成熟技术,相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(Microarray)或 DNA 芯片(又称为“基因芯片”)上,用于分析组织中基因的差异表达分析和基因诊断。用 CLG 蛋白特异的引物进行 RNA-聚合酶链反应(RT-PCR)体外扩增也可检测 CLG 蛋白的转录产物。
25

检测 CLG 蛋白基因的突变也可用于诊断 CLG 蛋白相关的疾病。CLG 蛋白突变的形式包括与正常野生型 CLG 蛋白 DNA 序列相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如 Southern 印迹法、DNA 序列分析、PCR 和原位杂交检测突变。另外,突变有可能影响蛋白的表达,因此用 Northern 印迹法、Western 印迹法可间接判断基因有无突变。
30

本发明的序列对染色体鉴定也是有价值的。该序列会特异性地针对某条人染色体具体位置且并可以与其杂交。目前,需要鉴定染色体上的各基因的具体位点。现在,只有很少的基于实际序列数据(重复多态性)的染色体标记物可用于标记染色体位置。根据本发明,为了将这些序列与疾病相关基因相关联,其重要的第一步就是将这些 DNA 序列
35 定位于染色体上。

简而言之,根据 cDNA 制备 PCR 引物(优选 15-35bp),可以将序列定位于染色体上。然后,将这些引物用于 PCR 筛选含各条人染色体的体细胞杂合细胞。只有那些含有相应于引物的人基因的杂合细胞会产生扩增的片段。

体细胞杂合细胞的 PCR 定位法,是将 DNA 定位到具体染色体的快捷方法。使用本发

- 盒(Pharmacia 公司)提取 mRNA。用 pCMV-script TMXR cDNA 文库构建试剂盒(Seratagene 公司)构建上述 mRNA 的 cDNA 文库。其中反转录酶改用 MMLV-RT-Superscript II(GIBCO BRL), 反转录反应在 42℃ 进行。转化 XL 10-Gold 感受细胞, 获得了 1×10^6 cfu/ μ g cDNA 滴度的 cDNA 文库。第一轮随机挑取 cDNA 克隆, 其后以高丰度 cDNA 克隆和已证明有抑
- 5 癌细胞生长功能的 cDNA 克隆为探针, 杂交筛选 cDNA 文库, 挑取弱阳性及阴性克隆。用 Qiagen 96 孔板质粒抽提试剂盒, 按厂家说明书进行质粒 DNA 的提取。质粒 DNA 和空载体同时转染肝癌细胞系 7721。100ng DNA 酒精沉淀干燥后, 加 6 μ l H₂O 溶解, 待转染。每份 DNA 样品中加 0.74 μ l 脂质体及 9.3 μ l 无血清培液, 混匀后, 室温放置 10 分钟。每管中加 150 μ l 无血清培液, 均分加入 3 孔生长于 96 孔板的 7721 细胞中, 37℃
- 10 放置 2 小时, 每孔再加 50 μ l 无血清培液, 37℃ 24 小时。每孔换 100 μ l 全培液, 37℃ 24 小时, 换含 G418 的全培液 100 μ l, 37℃ 24-48 小时, 边观察, 边换 G418 浓度不等的培液。约 2-3 次后, 直到镜检细胞有克隆形成, 计数。发现 PP3898(即 CLG)有抑制细胞克隆形成作用(抑制率在 50%或 50%以上), 结果如下表所示。

表 1 cDNA 克隆转染细胞(7721)克隆形成情况

cDNA 克隆名称	CDNA 克隆数(三个重复)	空载体克隆数(三个重复)
PP3898(即 CLG)	2 0 0	33 34 38

15

对 cDNA 克隆采用双脱氧终止法, 在 ABI377 DNA 自动测序仪上测定其一端近 500bp 的核苷酸序列。分析后, 确定为新基因克隆, 再进行完全测序, 发现 CLG 片段为非全长序列, 需用 RACE 方法获取全长 cDNA 克隆。

20 实施例 2: RACE 法获得全长 cDNA 克隆和 RT-PCR 法获得 CLC 基因

1. RACE 法获得全长 cDNA 克隆

对 PP3898 cDNA 克隆序列分析后发现基因尚不完整, 采用 Clontech 公司 SMART RACE cDNA 扩增试剂盒(Cat. No. K1811-1), 设计基因特异引物(如下表 2 所示), 按说明书进行操作, 获得全长克隆。

25

表 2 CLG 基因特异引物

克隆名称	特异引物 1(pp3898-NB)	特异引物 2(pp3898-B)
PP3898	5' TCATCCAGCCGGTCACTTGACTTGA 3'	5' GCCACAGCTGGTAGTTGGACTTGCC 3'

具体而言, 对 PP3898 克隆使用如下引物:

- 通用引物 mix(UPM) Long 5' CTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGT3'
- 巢式通用引物(NUP) 5' AAGCAGTGGTAACAACGCAGAGT 3'
- 30 pp3898-NB 5' TCATCCAGCCGGTCACTTGACTTGA 3'
- pp3898-B 5' GCCACAGCTGGTAGTTGGACTTGCC 3'

用人胎盘组织 mRNA 为起始材料, 按 Clontech 公司 SMART RACE cDNA 扩增试剂盒(Cat#1811-1)说明书获得 cDNA. 然后分别以 UPM 引物和基因特异的 pp3898-B 引物进行第一轮 PCR, 再以 NUP 引物和基因特异的 pp3898-NB 引物进行第二轮 PCR, 获得基因片

35 段。

反应条件如下: 94℃ 1 分钟, 一个循环; 94℃ 30 秒, 72℃ 4 分钟, 5 个循环;

GTCACCCCG AGGAGATCCA GCTGGGCGAG GACGAGGACG AGGACGAGAT GGACCTGGAG 2520
 CCCAACGAGG TTCGGCTGGA GCAGCAGAGC GTGCCAGCCG CAGTGTGTTGG GAGCCTGAAG 2580
 GAAGACTGAC CCGTCCCTCC CCCCTCCCA CCCCTCCCC AATACAGCTA CGTTGTGAAA 2640
 AAAAAAAAA AAAAAAAAA 2659

5

B: 氨基酸序列 (SEQ ID NO: 2) 长度: 855

1 MVVMARLSRP ERPDLVFEE DLPYEEI MR NQFSVKCWL R YIEFKQGAPK
 51 PRLNQLYERA LKLLPCSYKL WYRYLKARRA QVKHRCVTD P AYEDVNNCHE
 101 RAFVFMHKMP RLWLDYCF L MDQGRVTHTR RTFDRALRAL PITQHSRIWP
 151 LYLRFLRSH P LPETA VRGYR RFLKLSPE SA EEYIEYLKSS DRLDEAAQRL
 201 ATVVNDERFV SKAGKSNYQL WHELCDLISQ NPDQVQSLN V DAIIRGGLTR
 251 FTDQLGKLWC SLADYYIRSG HFEKARDVYE EAI RTVMTVR DFTQVFD SYA
 301 QFEESMIAAK METASELGRE EEDVDLELR LARFEQLISR RPLLLNSVLL
 351 RQNP HHVHEW HKRVALHQGR PREIINTYTE AVQTVDPFKA TGKPHTLWVA
 15 401 FAKFYEDNGQ LDDARVILEK ATKVNFKQVD DLASVWCQC G ELELRHENYD
 451 EALRLRLKAT ALPARRAEYF DGSEPVQNRV YKSLKVWSML ADLEESLGTF
 501 QSTKAVYDRI LDLRIATPQI VINYAMFLEE HKYFEESFKA YERGISLFWK
 551 PNVSDIWSTY LTKFIARYGG RKLERARDLF EQALDGCPPK YAKTLYLLYA
 601 QLEEEWGLAR HAMAVYERAT RAVEPAQQYD MFNIYIKRAA EIYGVTHTRG
 20 651 IQQKATIEVL S DEHAREMCLR FADMECKLGE IDRARIYSF CSQICDPRTT
 701 GAFWQTWKDF EVRHGNETI KEMLRIRRSV QATYNTQVNF MASQMLKVSG
 751 SATGTVSDLA PGQSGMDDMK LLEQRAEQLA AEAERDQPLR AQSKILFVRS
 801 DASREELAE L AQQVNPEEI Q LGEDEDEDEM DLEPNEVRLE QQSVPAAVFG
 851 SLKED

25

C. 核苷酸及氨基酸组合序列: (SEQ ID NO: 3) 克隆号: CLG(即 PP3898)

起始编码子: 22 ATG 终止编码子: 2589 TGA 蛋白质分子量: 100004.44

1 GGT ACC TGG GCA TCC AGA AAA ATG GTG GTG ATG GCG CGA CTC TCG CGG 48
 1 Met Val Val Met Ala Arg Leu Ser Arg 9
 30 49 CCC GAG CGG CCG GAC CTT GTC TTC GAG GAA GAG GAC CTC CCC TAT GAG 96
 10 Pro Glu Arg Pro Asp Leu Val Phe Glu Glu Asp Leu Pro Tyr Glu 25
 97 GAG GAA ATC ATG CGG AAC CAA TTC TCT GTC AAA TGC TGG CTT CGC TAC 144
 35 26 Glu Glu Ile Met Arg Asn Gln Phe Ser Val Lys Cys Trp Leu Arg Tyr 41
 145 ATC GAG TTC AAA CAG GGC GCC CCG AAG CCC AGG CTC AAT CAG CTA TAC 192
 42 Ile Glu Phe Lys Gln Gly Ala Pro Lys Pro Arg Leu Asn Gln Leu Tyr 57
 40 193 GAG CGG GCA CTC AAG CTG CTG CCC TGC AGC TAC AAA CTC TGG TAC CGA 240
 58 Glu Arg Ala Leu Lys Leu Leu Pro Cys Ser Tyr Lys Leu Trp Tyr Arg 73
 241 TAC CTG AAG GCG CGT CGG GCA CAG GTG AAG CAT CGC TGT GTG ACC GAC 288
 45 74 Tyr Leu Lys Ala Arg Arg Ala Gln Val Lys His Arg Cys Val Thr Asp 89
 289 CCT GCC TAT GAA GAT GTC AAC AAC TGT CAT GAG AGG GCC TTT GTG TTC 336
 90 Pro Ala Tyr Glu Asp Val Asn Asn Cys His Glu Arg Ala Phe Val Phe 105
 337 ATG CAC AAG ATG CCT CGT CTG TGG CTA GAT TAC TGC CAG TTC CTC ATG 384
 50 106 Met His Lys Met Pro Arg Leu Trp Leu Asp Tyr Cys Gln Phe Leu Met 121
 385 GAC CAG GGG CGC GTC ACA CAC ACC CGC CGC ACC TTC GAC CGT GCC CTC 432
 122 Asp Gln Gly Arg Val Thr His Thr Arg Arg Thr Phe Asp Arg Ala Leu 137
 55 433 CGG GCA CTG CCC ATC ACG CAG CAC TCT CGA ATT TGG CCC CTG TAT CTG 480
 138 Arg Ala Leu Pro Ile Thr Gln His Ser Arg Ile Trp Pro Leu Tyr Leu 153
 481 CGC TTC CTG CGC TCA CAC CCA CTG CCT GAG ACA GCT GTG CGA GGC TAT 528
 60 154 Arg Phe Leu Arg Ser His Pro Leu Pro Glu Thr Ala Val Arg Gly Tyr 169
 529 CGG CGC TTC CTC AAG CTG AGT CCT GAG AGT GCA GAG GAG TAC ATT GAG 576
 170 Arg Arg Phe Leu Lys Leu Ser Pro Glu Ser Ala Glu Glu Tyr Ile Glu 185
 577 TAC CTC AAG TCA AGT GAC CGG CTG GAT GAG GCC GCC CAG CGC CTG GCC 624
 65 186 Tyr Leu Lys Ser Ser Asp Arg Leu Asp Glu Ala Ala Gln Arg Leu Ala 201

	586	Gly Cys Pro Pro Lys Tyr Ala Lys Thr Leu Tyr Leu Leu Tyr Ala Gln	601
	1825	CTG GAG GAG GAG TGG GGC CTG GCC CGG CAT GCC ATG GCC GTG TAC GAG	1872
5	602	Leu Glu Glu Glu Trp Gly Leu Ala Arg His Ala Met Ala Val Tyr Glu	617
	1873	CGT GCC ACC AGG GCC GTG GAG CCC GCC CAG CAG TAT GAC ATG TTC AAC	1920
	618	Arg Ala Thr Arg Ala Val Glu Pro Ala Gln Gln Tyr Asp Met Phe Asn	633
	1921	ATC TAC ATC AAG CGG GCG GCC GAG ATC TAT GGG GTC ACC CAC ACC CGC	1968
10	634	Ile Tyr Ile Lys Arg Ala Ala Glu Ile Tyr Gly Val Thr His Thr Arg	649
	1969	GGC ATC TAC CAG AAG GCC ATT GAG GTG CTG TCG GAC GAG CAC GCG CGT	2016
	650	Gly Ile Tyr Gln Lys Ala Ile Glu Val Leu Ser Asp Glu His Ala Arg	665
15	2017	GAG ATG TGC CTG CGG TTT GCA GAC ATG GAG TGC AAG CTC GGG GAG ATT	2064
	666	Glu Met Cys Leu Arg Phe Ala Asp Met Glu Cys Lys Leu Gly Glu Ile	681
	2065	GAC CGC GCC CGG GCC ATC TAC AGC TTC TGC TCC CAG ATC TGT GAC CCC	2112
20	682	Asp Arg Ala Arg Ala Ile Tyr Ser Phe Cys Ser Gln Ile Cys Asp Pro	697
	2113	CGG ACG ACC GGC GCG TTC TGG CAG ACG TGG AAG GAC TTT GAG GTC CGG	2160
	698	Arg Thr Thr Gly Ala Phe Trp Gln Thr Trp Lys Asp Phe Glu Val Arg	713
	2161	CAT GGC AAT GAG GAC ACC ATC AAG GAA ATG CTG CGT ATC CGG CGC AGC	2208
25	714	His Gly Asn Glu Asp Thr Ile Lys Glu Met Leu Arg Ile Arg Arg Ser	729
	2209	GTG CAG GCC ACG TAC AAC ACG CAG GTC AAC TTC ATG GCC TCG CAG ATG	2256
	730	Val Gln Ala Thr Tyr Asn Thr Gln Val Asn Phe Met Ala Ser Gln Met	745
30	2257	CTC AAG GTC TCG GGC AGT GCC ACG GGC ACC GTG TCT GAC CTG GCC CCT	2304
	746	Leu Lys Val Ser Gly Ser Ala Thr Gly Thr Val Ser Asp Leu Ala Pro	761
	2305	GGG CAG AGT GGC ATG GAC GAC ATG AAG CTG CTG GAA CAG CGG GCA GAG	2352
35	762	Gly Gln Ser Gly Met Asp Asp Met Lys Leu Leu Glu Gln Arg Ala Glu	777
	2353	CAG CTG GCG GCT GAG GCG GAG CGT GAC CAG CCC TTG CGC GCC CAG AGC	2400
	778	Gln Leu Ala Ala Glu Ala Glu Arg Asp Gln Pro Leu Arg Ala Gln Ser	793
	2401	AAG ATC CTG TTC GTG AGG AGT GAC GCC TCC CGG GAG GAG CTG GCA GAG	2448
40	794	Lys Ile Leu Phe Val Arg Ser Asp Ala Ser Arg Glu Glu Leu Ala Glu	809
	2449	CTG GCA CAG CAG GTC AAC CCC GAG GAG ATC CAG CTG GGC GAG GAC GAG	2496
	810	Leu Ala Gln Gln Val Asn Pro Glu Glu Ile Gln Leu Gly Glu Asp Glu	825
45	2497	GAC GAG GAC GAG ATG GAC CTG GAG CCC AAC GAG GTT CGG CTG GAG CAG	2544
	826	Asp Glu Asp Glu Met Asp Leu Glu Pro Asn Glu Val Arg Leu Glu Gln	841
	2545	CAG AGC GTG CCA GCC GCA GTG TTT GGG AGC CTG AAG GAA GAC TGA CCC	2592
50	842	Gln Ser Val Pro Ala Ala Val Phe Gly Ser Leu Lys Glu Asp ***	856
	2593	GTC CCT CCC CCC TCC CCA CCC CCT CCC CAA TAC AGC TAC GTT TGT AAA	2640
	2641	AAA AAA AAA AAA AAA AAA A	2659

实施例 4: 同源比较和结构分析

- 55 应用 Blast 软件, 对 CLG 的氨基酸序列进行同源性分析, 发现与果蝇、线虫和酵母的细胞周期调控蛋白高度同源: 与果蝇的细胞周期调控蛋白(crooked neck, crn)有 41%(157/375)的同源性, 与线虫的细胞周期调控蛋白具有 66%(557/837)的同源性, 与酵母的细胞周期控制蛋白有 37%(228/599)的同源性。

60 Query = CLG Sbjct = crn (果蝇)
 > crn 基因 长度 = 702
 分值 = 69.9 bits (168), 预计值 = 6e-11
 相同性 = 88/375 (23%), 相似性 = 157/375 (41%), 缺口 = 33/375 (8%)

- Query: 498 GTFQSTKAVYDRILDLRIATPQIVINYAMFLEEHKYFEESFKAYERGISLFKWPVNSDIW 557
GT +S + VYD+++LR+A+PQ+++NYAMFLEE++YFE +F+AYE+GI+LFKWP V DIW
Sbjct: 502 GTVESCRKVYDKMIELRVASPMIMNYAMFLEENEYFELAFQAYEKGIALFKWPGVFDIW 561
- 5 Query: 558 STYLTKFIARYGGRKLERARDLFEQALDGCPPKYAKTLYLLYAQLEEEWGLARHAMAVYE 617
+TYL KFI RYGG+KLERARDLFEQ L+ CPP +AK ++LLYA+LEEE GLARHA+++Y
Sbjct: 562 NTYLVKFIKRYGGKKLERARDLFEQCLENCPPTHAKYIFLLYAKLEEEHGLARHALSIYN 621
- 10 Query: 618 RATRAVEPAQQYDMFNIIYIKRAAEIYGVTHTRGIYQKAIEVLSDEHAREMCLRFADMECK 677
RA V+ A + M+NIYIK+ E+YG+ R I+++AI L ++ +R M LR+A +E
Sbjct: 622 RACSGVDRADMHSYNIYIKKVQEMYGIAQCRPIFERAISELPEDKSRAMSLRYAQLLETT 681
- 15 Query: 678 LGEIDRARAIYSFCSQICDPRTTGAFWQTKDFEVRHGNETIKEMLRIRRSVQATYNTQ 737
+GEIDRARAIY+ ++I DP+ FW TWK+FEV HGNE T+++MLR+RRSV+A+YN
Sbjct: 682 VGEIDRARAIYAHAAEISDPKVHVKFWDTWKNFEVAHGNEATVRDMLRVRRSVEASYNVN 741
- 20 Query: 738 VNFMASQM-LKVSGSATGTVSDLAPGQSGMDMKXXXXXXXXXXXXDQPLRAQSKIL 796
V + QM + A T + P S +D + Q + I
Sbjct: 742 VTLTSVQMRVDAERKAQETTTSSNPMDLSDQDQDQPSDGCAGSIT---QVSMNKGNI 795
- Query: 797 FVRSDASREELAEALQQVNPEEIQLGXXXXXXXXXXXXXVRLEQQSVPAVFGSLK 853
FVR + + NP+EI L + + + VPA +FG+LK
Sbjct: 796 FVRGAG---KTVQQTENPDEIDLDEDDDDDEDDGGDADISV---KVVPAQIFGNLK 847
- 25 Query = CLG Sbjct = 推定的细胞周期调控蛋白 (酵母)
> 推定的细胞周期调控蛋白 长度 = 674
分值 = 78.4 bits (190), 预计值 = 1e-13
相同性 = 125/599 (20%), 相似性 = 228/599 (37%), 缺口 = 115/599 (19%)
- 30 Query: 24 YEEEIMRNQFSVKWLRYIEFKQGAPK-PRLNQLYERALKLLPCSYKLWYRYLKARRAQV 82
+E+ I RN+ ++ W+RY +++ + R ++ERAL + LW +Y++ ++
Sbjct: 59 FEDAIRNRLAMGHWMRYGQWELDQKEFARARSVFERALDVDSTYIPLWLKYIEC---EM 115
- 35 Query: 83 KHRCTDPAYEDVNNCHERAFVFMHKMPRLWLDYQFLMDQGRVTHTRRTFDRLRALPI 142
K+R + N +RA + ++ +LW Y G +T R+ F+R L+ P
Sbjct: 116 KNRNINH---ARNLFDRAVTQLPRVDKLWYKYVYMEMLGNITGCRQVFERWLKWEF- 169
- 40 Query: 143 TQHSRIWPLYRLRSHPLPETAVRGYRRFLKLSPEAAEYIEYLKSSDRLEAAQRLAT 202
W Y+R R + E A Y RF+ + PE ++ + + + AA
Sbjct: 170 --DENCWMSYIRMERRYHENERARGIYERFVVHPE-VTNWLRWARFEEECGNAA----- 221
- 45 Query: 203 VVNDERFVSKAGKSNYQLWHELCDLISQNPDKVQSLNVDIAIRGGLTRFTDQGLKWLCSL 262
N +V +DA+ + L + + +
Sbjct: 222 ---NVRQVYLAIDALGQEFLE---RFFIAF 247
- 50 Query: 263 ADYYIRSGHFEKARDVYEEAIRVTMTVRDFTQVFDSYAQFEESMIAAKMETASXXXXXX 322
A + IR +E+AR +++ AI M +++ Y FE+
Sbjct: 248 AKFEIRQKEYERARTIFKYAI-DFMPRSKSMELYKEYTHFEKQF----- 290
- 55 Query: 323 XXXXXXXXXXXFEQLISRRPLLNSVLLRQNPHHVHEWHKRVALHQ-GRPREIINTYTE 380
E + + L LL+ +P+ W + L + G I TY +
Sbjct: 291 ---GDHLGVSTVLDKRRQLQYEKLLKDSPPYDYDTWLDLLKLEESAGDINTIRETYEK 344
- 60 Query: 381 AVQTVDPF---KATGKPHTLWVAFKYE-DNGQLDDARVILEKATKVNFKQVDDLASVW 436
A+ V A + +W+ + F E D +D AR + +A K+ + A +W
Sbjct: 345 AIAKVPVVEKNAWRRYVYIWLNYCLFEEIDVKDVRARKVYQEALKLIPHKFTFAKLW 404
- 65 Query: 437 CQCGELELRHENYDEALRLRKATAL---PARRAEYFDGSEPQN---RVY----- 481
ELR D A + L +A + P Y + + ++ R+
Sbjct: 405 LMYAMFELRQRKIDVARKTLGRALGMCPKPLFRGYIEFEDAIFDRCRILYEKWIYD 464
- Query: 482 -KSLKVVSMADLEESLGTFSQSTKAVYDRILDLRI-ATPQIVIN-YAMFLEEHKYFEESF 538
++ W A LE LG +A+Y+ ++ I TP++V Y F E + ++
Sbjct: 465 PEACAPWLGAALETCLGDSDRARALYNLAVNQPILETPELVWKAYIDFEFEEMEYGGAR 524

实施例 6: CLG 对人肝癌细胞集落形成 (Colony Formation) 的抑制作用

RT-PCR 获得的含完整编码区的 CLG 基因 (实施例 2), 采用 AdvanTage™ PCR Cloning 试剂盒 (Clontech, Cat#: K1901-1) 克隆于 pT-Adv 载体 (Clontech), 然后 EcoR I 酶切, 回收含完整编码区的 CLG 基因片段, 再亚克隆至真核细胞表达载体 pCMV-Script (Stratagene, Cat#: 212220), 获得质粒 pCMV-Script/CLG, 用 Qiagen 质粒抽提试剂盒进行质粒 DNA 的提取, 供转染用。

取 5 μ g CLG 质粒 DNA (载体为 pCMV-Script), 加无血清无抗生素的 DMEM 到 600 μ l, 空白对照用消毒 milli-Q 水代替质粒 DNA。另取 25 μ l 脂质体 Lipofectamine (GIBCO BRL), 加无血清无抗生素的 DMEM 至 600 μ l, 将上述质粒 DNA 和脂质体混合, 轻柔摇匀, 室温放置 30-45 分钟, 再加入无血清无抗生素的 DMEM 1800 μ l, 总体积为 3ml。将 3ml 转染复合物加入生长于 6cm 培养皿的 SMMC-7721 细胞中 ($1-2 \times 10^5$ 细胞 / 每培养皿), 37 $^{\circ}$ C 培养 6 小时后, 换用全培养液, 24 小时后, 再换用 G418 的全培养液, 约两周左右出现集落, 结晶紫 (crystal violet) 染色, 观察着色的集落数目。

结果显示, CLG 全长 cDNA 转染 SMMC-7721 细胞能显著抑制集落形成。转染 CLG 后 SMMC-7721 细胞只形成 3 个集落, 而转染空载体的 SMMC-7721 细胞则形成 56 个集落, CLG 基因对人肝癌细胞 SMMC-7721 在体外有明显的抑制作用 (图 2A 和 2B)。

实施例 7: CLG 基因转化细胞的成瘤实验及凋亡检测

通过实施例 6 相同的方法, 用 CLG 基因转染人肝癌细胞 SMMC-7721, 然后将转化细胞及未转染 CLG 基因的对照组细胞扩大培养, 皮下接种裸鼠, 观察 CLG 基因对成瘤作用的影响, 实验分为 CLG 转染组和对照组, 每组 6 只裸鼠, 每只裸鼠接种 2×10^6 细胞, 观察期为 6 周, 成瘤实验结果见表 3。

表 3 SMMC-7721 细胞成瘤实验结果

组 别	瘤 重 (g)						平均瘤重 (g)	T 检验结果
	1	2	3	4	5	6		
对照	0.24	0.11	0.15	0.16	0.12	0.18	0.16	P<0.05
CLG	0.14	0.07	0.01	0.03	0.10	0.07	0.07	

25

在成瘤实验中, 6 只实验裸鼠接种 CLG 转染的 SMMC-7721 细胞后, 形成的肿瘤平均瘤重为 0.07 克, 而 6 只接种 SMMC-7721 细胞自身对照组所形成的瘤重为 0.16 克, 经平均值的成对两样本分析 T 检验, 二者间有显著性差异, $p < 0.05$, 抑瘤率为 50%。

CLG 转染的 SMMC-7721 细胞, 接种裸鼠后形成的肿瘤组织, 作石蜡切片, 用 TUNEL 法进行凋亡原位检测 (Roche 公司), 可见蓝紫色的阳性凋亡细胞, 其中在坏死灶内可见较多阳性凋亡细胞 (图 3A 所示), 在未坏死区可见散在的凋亡细胞 (图 3B 所示)。

-80 $^{\circ}$ C 冷冻的成瘤组织碾碎后, 用含 RNA 酶和蛋白酶 K 的 DNA 抽提溶液消化, 50 $^{\circ}$ C 消化 2 小时, 然后酚抽提 2 次, 酒精沉淀 DNA, 1.5% 琼脂糖凝胶电泳分析, 观察是否有“梯形” DNA 条带。转染 CLG 全长基因的 SMMC-7721 细胞形成的肿瘤组织 DNA 经 1.5% 琼

权 利 要 求 书

- 1.一种分离的人CLG蛋白,其特征在于,它包括:具有SEQ ID NO: 2氨基酸序列的多肽;或其保守性变异多肽、或其活性片段、或其活性衍生物。
- 5 2.如权利要求1所述的多肽,其特征在于,该多肽包含SEQ ID NO: 2的氨基酸序列。
- 3.一种分离的多核苷酸,其特征在于,它包含一核苷酸序列,该核苷酸序列与选自下组的一种核苷酸序列有至少85%相同性:
 - (a)编码如权利要求1和2所述多肽的多核苷酸;
 - (b)与多核苷酸(a)互补的多核苷酸。
- 10 4.如权利要求3所述的多核苷酸,其特征在于,该多核苷酸编码的多肽具有SEQ ID NO:2的氨基酸序列。
- 5.如权利要求3所述的多核苷酸,其特征在于,该多核苷酸的序列选自下组:
SEQ ID NO:3中第22—2258位的编码区序列,或1—2659位全长序列。
- 6.一种载体,其特征在于,它含有权利要求3所述的多核苷酸。
- 15 7.一种遗传工程化的宿主细胞,其特征在于,它是选自下组的一种宿主细胞:
 - (a)用权利要求6所述的载体转化或转导的宿主细胞;
 - (b)用权利要求3所述的多核苷酸转化或转导的宿主细胞。
- 8.一种具有人CLG蛋白活性的多肽的制备方法,其特征在于,该方法包含:
 - (a)在适合表达人CLG蛋白的条件下,培养权利要求7所述的宿主细胞;
 - (b)从培养物中分离出具有人CLG蛋白活性的多肽。
- 20 9.一种能与权利要求1所述的人CLG蛋白特异性结合的抗体。
10. 一种药物组合物,其特征在于,它含有安全有效量的权利要求1所述的多肽以及药学上可接受的载体。

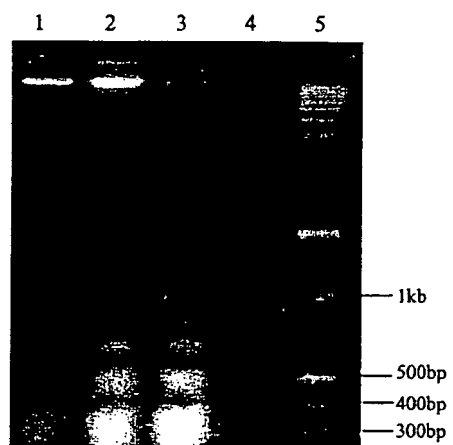


图 4

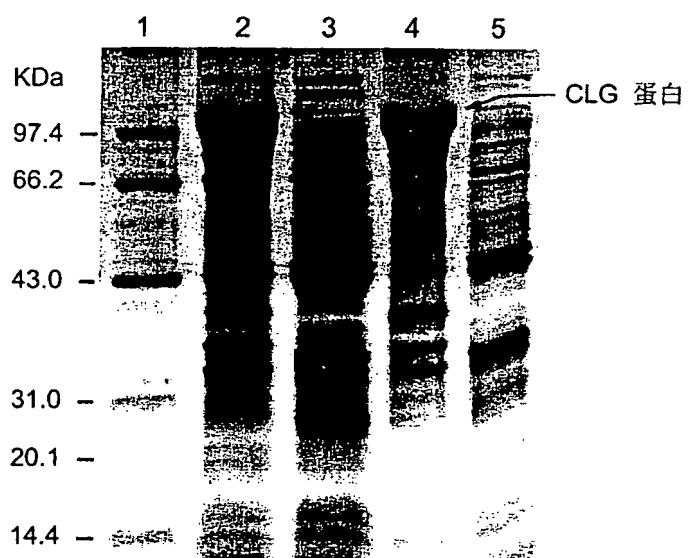


图 5

国际检索报告

国际申请号

PCT/CN01/00121

A. 主题的分类 IPC ⁷ C07K14/47, C12N15/12, C12N15/63 按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类		
B. 检索领域 检索的最低限度文献(标明分类体系和分类号) IPC ⁷ C07K14/47, C12N15/12, C12N15/63 包含在检索领域中的除最低限度文献以外的检索文献 中国专利文献数据库, 中文科技期刊数据库 在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词) GenBank, EPOQUE, BA, MEDLINE		
C. 相关文件		
类 型*	引用文件, 必要时, 包括相关段落的说明	相关的权利要求编号
A	RNA, 1999, 5(8):1042-54 , 见摘要	1 - 10
A	Genes Dev, 1991, 5(6):1080-91 见摘要	1 - 10
<input type="checkbox"/> 其余文件在 C 栏的续页中列出。 <input type="checkbox"/> 见同族专利附件。		
<p>* 引用文件的专用类型: “A” 明确表示了一般现有技术、不认为是特别相关的文件 “E” 在先文件, 但是在国际申请日的同一日或之后公布的 “L” 对优先权要求可能产生怀疑或者用来确定另一篇引用文件的公布日期或其它特殊理由而引用的文件(如详细说明) “O” 涉及口头公开、使用、展览或其他手段的文件 “P” 在国际申请日之前但迟于所要求的优先权日公布的文件 “T” 在国际申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理 “X” 特别相关的文件: 当该文件被单独使用时, 要求保护的发明不能认为是新颖的或不能认为具有创造性 “Y” 特别相关的文件: 当该文件与其他一篇或多篇这类文件结合在一起, 这种结合对本领域技术人员是显而易见的, 要求保护的发明不能认为具有创造性 “&” 同族专利成员的文件</p>		
国际检索实际完成的日期 12. 3 月 2001(12.03.01)		国际检索报告邮寄日期 22. 3 月 2001(22.03.01)
国际检索单位名称和邮寄地址 中国专利局 中国北京市海淀区西土城路 6 号(100088) 传真号: 86-010-62019451		授权官员 孙广秀 电话号码: 86-010-62093884